Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPACT OF BROKEN LINKS AND DEAD CODES ON OPEN-SOURCE REPOSITORIES: AN AI AUTO ENCODER APPROACH FOR SENSITIVE DATA PROTECTION

M. MUTHALAGU[1], DR. R. RATHINASABAPATHY[2]

[1] Research scholar, Department of Computer Application, Madurai Kamaraj University, Madurai.
[2] Associate Professor, Department of Computer Application, Madurai Kamaraj University, Madurai.
E-mail: [1]muthalagucs76@gmail.com, [2]saba.dca@mkuniversity.ac.in

ABSTRACT

This article examines the effects of broken links and obsolete code on open-source repositories' usefulness, security, and sustainability. Defunct links, sometimes resulting from outdated or deleted external resources, hinder developers' access to essential documentation, libraries, and tools. Dead code, denoting abandoned or old code inside repositories, complicates maintenance and exposes possible risks. Furthermore, the proliferation of big data introduces distinct issues in managing large volumes of unstructured sensitive information, especially regarding extraction and analysis. This paper suggests implementing a Deep Autoencoder (DAE) model for protecting sensitive data. This method utilizes AI-driven auto encoders to identify sensitive data patterns, encrypt them, and discover weak connections and obsolete code for effective elimination. The optimized DAE algorithm demonstrates enhanced performance with increased detection rates, diminished false positives, and minimized failure jitter, making it a reliable option for risk assessment and prolonging the lifespan of open-source repositories. The results underscore the need for consistent maintenance and community cooperation to enhance open-source software ecosystems' quality, dependability, and security.

Keywords: *Data Mining, Sensitive Data, DAE (Deep Autoencoder), Broken Links, Dead Code, Open-Source Repositories, Code Maintenance, Software Sustainability, Code Quality, Developer Collaboration, Repository Integrity.*

1. INTRODUCTION

This guide provides details to assist authors in understanding that systemic financial risk denotes possible hazards that may affect the whole monetary system, exemplified by the global financial crisis. Since the twentieth century, the frequency and severity of financial crises have escalated. Notwithstanding significant enhancements in the worldwide financial landscape, the international financial sector continues to recover from the ramifications of previous crises. The interdependence of global economies, propelled by globalization, has increased the probability and consequences of financial risks. This interconnectivity implies that disruptions in one location may rapidly propagate to another, rendering financial systems susceptible. China has established itself as a pioneer in financial innovation, especially in digital finance, by developing new financial systems. In recent years, alternative payment methods have supplanted conventional ones, and there have been notable developments in digital insurance, intelligent investing, and online finance. Nonetheless, the digital financial system is becoming more vulnerable to internal disturbances. The swift propagation of dangers over the Internet may rapidly transform these difficulties into systemic

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

risks, jeopardizing enterprises and financial stability.

Conventional financial risk management methods have shown insufficient forecasting and alleviation of these risks. The failure to adequately foresee dangers arises from the absence of prompt and effective risk indicators. Conventional risk warning models depend on statistical data, although they often do not provide dependable early warnings. As financial systems become more intricate, scholars and practitioners concur that the attributes of risk models are essential for their effective implementation. The emergence of big data has generated extensive unstructured data, which might enhance financial risk forecasting. This data, accessible in many forms, may be analyzed using artificial intelligence (AI) to detect critical risk indicators effectively. Advancements in artificial intelligence regarding natural language processing (NLP), picture identification, and cognitive perception have created new opportunities for obtaining essential information from unusual sources, including scanned documents, photos, and textual data.

Optical Character Recognition (OCR) and Natural Language Processing (NLP) may transform unstructured data into actionable insights. For instance, OCR techniques may retrieve financial data from unconventional documents, whilst NLP can analyze real-time insights from news and public sentiment, uncovering financial hazards before their escalation. Incorporating these technologies into financial systems offers enhanced capabilities for risk prediction. Images and words, as novel data sources, pose issues owing to their varied forms, substantial volume, and frequency. In contrast to conventional data, gathered mainly by governments and organizations, picture and text data are multisource and heterogeneous, rendering their standardization and processing intricate. The velocity of processing unstructured data has become essential. Advancements in deep learning, particularly Deep Auto Encoders (DAE), have enabled the more efficient processing of enormous datasets, resulting in enhanced accuracy in financial risk assessments. These strategies reduce erroneous detections and improve the accuracy of risk estimation.

Moreover, open-source code repositories serve as collaborative platforms for developers to contribute to and exchange code. Over time, these repositories may gather broken links and obsolete code, resulting in software inefficiencies. Defunct links, including obsolete API references, may result in runtime issues and diminish the user experience. Dead code, including useless variables and functions, adds extra complexity, increasing the likelihood of mistakes during further modifications. This study dramatically advances financial risk management by solving difficulties with novel techniques, including integrating AI and big data analytics. Using open-source platforms for code-sharing, coupled with AI-generated insights from unstructured data, presents a novel advancement in financial risk forecasting, facilitating more precise, efficient, and dependable early warning systems in a progressively linked digital landscape.

2. RELATED WORK

2.1 Decision Making and Support System

Decision-making is an essential competency for efficient management inside any organization, and the intricacies of the contemporary corporate landscape amplify the need for systematic, data-informed methodologies. As organizations get more complex, decision-making becomes more difficult, leading to the heightened significance of Decision Support Systems (DSS) across diverse sectors. These systems aid managers in managing the intricacies of corporate operations by using data analysis and predictive modelling to provide insights that guide essential choices.

2.2 Quality Function Deployment System and Fuzzy Logic System

Yazdani et al. investigated the use of Quality Function Deployment (QFD) to determine optimum solutions inside the agricultural supply chain. Quality Function Deployment (QFD) is a systematic methodology for product development that facilitates the conversion of client needs into technical specifications, enhancing customer

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

satisfaction. This research emphasizes the need to incorporate systematic decision-making frameworks into supply chain management, a crucial element for strengthening agricultural operational efficiency. In the healthcare industry, Scalia et al. used a multi factorial assessment method for pancreatic islet transplantation, an auspicious therapy for diabetes. Their study showed how decision-making frameworks might enhance the success rates of intricate medical operations. Fuzzy systems facilitated decision-making under uncertain conditions, particularly in medical diagnostics. Fuzzy systems enable decision-makers to analyze ambiguous or indistinct information, which is especially beneficial when exact data is lacking.

2.3 Artificial Emotions and Investment Decision-Making Systems

Cobrerra-Paniagua et al. proposed the notion of artificial emotions to augment the autonomy of decision-making systems, especially within financial investments. This system was designed to replicate emotional reactions, often essential to human decision-making, enhancing the system's capacity for autonomous investment choices. Artificial emotions indecision-making systems signify an innovative method for emulating human decision-making processes. There has been an emphasis on creating adaptive stock index trading decision systems that use Multiple Criteria Decision Making (MCDM) methodologies. MCDM entails assessing several variables to create judgements that reconcile opposing goals, such as risk and return in financial trading. Researchers have suggested incorporating these technologies into Early Warning technologies (EWS), which use predictive algorithms to detect possible financial concerns before their escalation.

2.4 Neural Networks and Financial Expert Systems

Kim et al. used Artificial Neural Networks (ANN) as economic indicators in the Korean financial market, using Early Warning System (EWS) models to predict stock prices. The benefit of using neural networks lies in their capacity to learn and adjust to new data, rendering them exceptionally proficient in financial forecasting, where market circumstances constantly evolve. Using neural networks allowed the researchers to exploit trading opportunities while reducing losses. Weng et al. created a Financial Expert System (FES) that evaluated sentiment ratings from news stories to forecast short-term stock prices. This system used quantitative and qualitative sentiment analysis, illustrating how varied data sources may enhance decision-making in financial markets. The use of machine learning and natural language processing in this context underscores the growing significance of AI in financial decision-making systems.

2.5 Investor Insights in Risk Evaluation

Although several studies emphasize the company perspective, it is as vital to evaluate the investor's standpoint. Businesses and investors often prioritize risk assessment differently, and comprehending both viewpoints is essential for informed decision-making. Previous studies have suggested a technique that combines business risk with investor viewpoints to provide a more holistic understanding of possible dangers. This method enables organizations to formulate risk management plans corresponding to internal goals and external investor anticipations.

2.6 Preemptive Alert Mechanisms for Financial Risk

Early Warning Systems (EWS) are essential for detecting and alleviating financial threats. These systems use predictive algorithms to anticipate unusual circumstances impacting financial stability, such as market declines or economic crises. Traditional Early Warning System methodologies often use logistic and probit regression models to evaluate the probability of a financial crisis. Researchers have been striving to improve these models by integrating information relevant to specific nations or industries, yielding more precise forecasts. EWS models in the financial sector are designed to identify and disclose risks that may affect financial statements. These systems are designed to notify enterprises of possible threats before their occurrence, facilitating proactive measures. Earlier research mainly concentrated on business- centric assessments, often neglecting investors' viewpoints.

Incorporating both business and investor perspectives is crucial for developing comprehensive risk assessment systems that meet the requirements of all stakeholders.

2.7 Influence of Code Repositories on Decision-Making Systems

In software development, especially in open-source projects, the administration of code repositories is essential for sustaining system efficiency. Obsolete code and malfunctioning linkages are prevalent problems that may considerably impair the efficacy of decision-making systems. Dead code denotes obsolete code segments that persist in the code base, resulting in inflated and inefficient systems. Conversely, defective linkages may lead to runtime faults that impair the operation of decision-making systems. Dead code may hinder code optimization and extend compilation durations, whilst broken links can lead to significant mistakes in the operation of decision-making systems. To sustain superior software systems, it is essential to audit codebases and eliminate outdated or defective components routinely. Automated technologies like SonarQube and ESLint have been created to identify dead code and broken links, enhancing system performance and maintainability.

2.8 Detection and Prevention of Code Issues

Various techniques have been developed to identify and mitigate problems associated with dead code and broken links. Manual detection, while comprehensive, is laborious and susceptible to human mistakes, particularly in extensive codebases. Automated detection methods, including static analysis tools and coverage analysis instruments, provide a more scalable and consistent method for discovering these vulnerabilities. These solutions may be included in development environments and continuous integration/continuous deployment (CI/CD) pipelines, offering real-time feedback to developers. Hybrid approaches integrating automatic detection with human review provide a balanced strategy, ensuring the resolution of basic and complicated code errors. Consistent upkeep of code repositories is crucial for maintaining the efficiency, security, and reliability of decision-making systems. Research on decision-making systems emphasizes the growing significance of sophisticated approaches, like Quality Function Deployment (QFD), fuzzy systems, and artificial neural networks, in facilitating intricate decision-making processes. Incorporating investor viewpoints into risk evaluation and establishing Early Warning Systems illustrates the progressive nature of decision support systems in finance. In software development, the administration of code repositories is essential for preserving the efficacy and security of decision-making systems. Automated tools and hybrid detection approaches provide efficient solutions for resolving code errors, guaranteeing that decision support systems operate efficiently in intricate and dynamic situations.

3. METHODOLOGY

The research focused on detecting future crises primarily seeks to detect warning indicators via a comprehensive analysis of company-related data. Disclosure data and financial statements provide an impartial representation of the firm. Consequently, these two categories of information are often used to identify danger signals. These data facilitate the reconstruction of prior actions and events of firms about credit. They can aid us in assessing the health of firms and identifying those that need the most attention from us. They may assist us in pinpointing the firms that need our utmost attention. It may be feasible to ascertain a firm's reliability by analyzing financial accounts. Nonetheless, there may be undisclosed risks, including substantial occurrences or accounting fraud. Non-financial information, such as investor sentiment, may act as a leading signal of possible issues for these enterprises. Businesses use opinion mining to uncover previously hidden information or to assess customer sentiment. A risk warning signal particular to a company is generated by amalgamating these two distinct types of information into a single measure. The distinguishing element of the risk indicator is the potential for a credit event; hence, this component should be the primary focus of the indicator's attention. The computed indicator serves as an additional instrument for assessing the amount of risk to which the firm is exposed. The concluding phase of many business processes ensures that the organization can identify real-time risk indicators. This study examines a crisis from two unique viewpoints, as seen in Fig.1. This perspective encompasses corporate data, called opinion mining, which may objectively characterize the business's characteristics and stakeholders' attitudes. A comprehensive risk assessment may provide a more accurate determination of the company's susceptibility to risk.

3.1 Data Preprocessing

The raw data has deficiencies that must be addressed with suitable information before further investigation. This procedure eliminates columns with substantial data deficiencies and particular samples that lack integrity to address absent units. Each information segment correlates to distinct features and qualities, whereas significant data gaps are omitted from the study. Rows denote individual samples, and in the final analysis, samples with substantial missing data are penalized and regarded as inadequate predictors. Rows represent the samples.

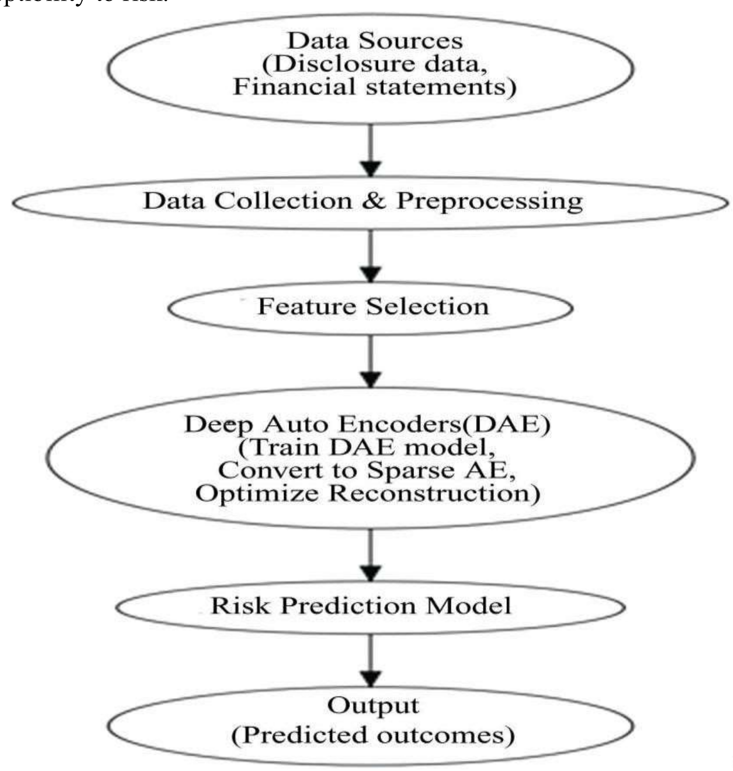

Figure1:Risk Detection in Financial Transactions

Vertical and lateral data cleaning are procedures used to guarantee that data may be efficiently employed in future analyses. The quantity of material excluded from consideration throughout these steps is negligible in the total data gathering. After concluding the data-cleaning procedure, we did not notice a substantial decrease in the sample size we had been using. A sufficient sampling was performed to guarantee that the final product would be fine with these minor modifications. This was accomplished by sampling an adequate amount of the finished product. During our data-gathering procedure, some categories still needed to be included, leading to the absence of assigned numerical values. We determined that the ratio of unused occurrences is sufficiently low to warrant eliminating NULL values from the dataset. This is the conclusion we have arrived at.

Data deletion is executed according to the number of rows and columns containing NULL values. This minimizes the danger of unintentionally obscuring any latent correlations within the data. Before picking a machine learning model, it is essential to identify the characteristics that are critical for constructing our final model.

3.2 Selection of Features

We have seen that the gathered material has considerable redundancy, necessitating its removal. Most of these factors will not affect the prediction's result, and some may even obfuscate the analysis. B Filtering is particularly essential for attributes that signify inconsequential details in isolation. One should employ mathematical analysis for a more compelling study than just assessing the importance of each element in isolation. This method accounts for many factors, resulting in a more thorough analysis.

The choice of characteristics for data analysis directly influences the model's accuracy. When the filtering procedure does not substantially impact model correctness, it is preferable to use fewer features, notably when the Pearson correlation coefficient method may diminish the required number of variables for the model. The Pearson correlation coefficient indicates the degree of association between two variables. To get the Pearson correlation coefficient, divide the correlation by the standard deviation, as seen in...

The Pearson correlation coefficient was designed to fulfil this need. This statistic assists in pinpointing the aspects that most significantly impact the investigation's conclusion with enhanced precision. Utilizing the Pearson correlation coefficient method may decrease the number of variables required for the model. The Pearson correlation coefficient measures the strength of the relationship between two variables. Divide the correlation by the standard deviation to get the Pearson correlation coefficient, as shown in Equation 1.

$$\rho x, y = \frac{Con(x,y)}{\underset{xy}{\sigma\sigma}} \quad (1)$$

On the other hand, it would be more intuitive to approach specific instances in the following way, as demonstrated in Equation (2).

$$r = \frac{N\sum i\, xiyi - \sum i\, xi \sum i\, yi}{\sqrt{N\sum(xi)2 - \sum(xi)2}\sqrt{N\sum(yi)2 - \sum(yi)2}} \quad (2)$$

The variable's value may vary from -1 to 1, with greater absolute values indicating a more vital link. We may modify the threshold value to regulate the number of most interrelated aspects, enabling us to discern closely correlated traits. The screening process has led to enhanced productivity and efficiency. All datasets have been purified, standardized, and prepared to guarantee precision in future analytical models. Data preprocessing is essential to ensure reliable and impartial results in subsequent analyses.

3.2 Deep Autoencoders

The purpose of neural networks, such as autoencoders, is to retain the input data so that it may be reconstructed as output data later. Using the trial set, it is essential to ascertain ways to extract the principal features from the input.

$$\{(1),(2),\ldots x(n)\},\text{where } x(1)\in R^d \quad (3)$$

Figure 2 illustrates the fundamental autoencoder, which includes a single input layer. One hidden layer and one output layer. The autoencoder model begins by encoding the singular input x(i) to the hidden layer y(x(i)), which is then decoded to produce the output layer z(x(i)), therefore indicating the success of the autoencoder model.

$$y(x)=f(W_1x+b) \quad (4)$$

$$z(x)=g(W_2x+c) \quad (5)$$

W_1 represents the weight matrix, and b is the encoded vector. W_2 represents the decode matrix, and c signifies the decoding vector. The logistic sigmoid function is represented as follows in Equation (6).

$$f(x).=1/(1+\exp(-x)) \quad (6)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

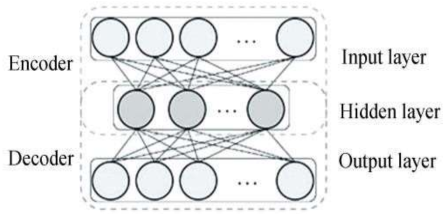

Figure2 Autoencoder Processes

The autoencoder concept necessitates an input layer represented by I and an encoding function defined by f. This combination finally results in creating an estimated output layer (y). The auto encoder model employs a decoder function(g) to reconstruct the original input layer(x), resulting in the development of the output layer (Z). The scaling with the loss function LH is (xz), which dictates the reconstruction error. The best configuration for the reconstruction parameters is achieved by minimizing the function$L$ $(X,)$ and determining the ideal settings for the reconstruction parameters.

$$\theta = arg\theta$$

$$N$$

$$minL(X,Z) = argmin0.5L(X,Z)\sum \| x(i) - z(x(i)) \|^2 \quad (7)$$

The dimension of the autoencoder model's hidden layer, often equal to or exceeding that of the output layer, is a considerable issue in this study field. The arrangement of the components of the model typically addresses this concern. The autoencoder model was converted into a sparse autoencoder using a nonlinear autoencoder with a hidden layer that exceeds the input layer by one unit in conjunction with the sparsity constraint approach. The sparsity constraint approach was used to achieve this objective effectively. We used a sparsity constraint to achieve a sparse representation and to reduce the error incurred during the reconstruction phase.

3.3 Core-Based DAE

The DAE and its parameter setup for congestion prediction may be assessed against two sophisticated deep-learning neural network models that function as benchmarks. This method enables the assessment of the DAE's efficacy and setup in forecasting congestion. We have implemented a system to make forecasts over three separate time horizons for comprehensive comparisons and assessments. The research evaluates the effectiveness of traffic congestion forecasts by using the Mean Absolute Error (MAE) and the weighted Mean Squared Error (WMSE). The current traffic congestion level (ctij), projected traffic congestion level (btij), and penalty weight (wtij) at time t are denoted by coordinates i and j in the D-matrix, which represents the highway transportation network. These positions on the grid represent the current congestion level. This matrix facilitates the visualization of information. If uneven congestion intensity distributions pose an issue, one method is to heuristically modify them to amplify penalties for erroneous forecasts of nonlinear congestion levels. The relevant acronyms for these parameters denote the grid's width (W) and height (H).

$$MAE = \frac{1}{W\ X\ H} \sum_{i=1}^{W} Hj \left\| ct_{ij} - ct_{ij} \right\| ..1 \quad (8)$$

$$MAE = \frac{1}{W\ X\ H} \sum_{i=1}^{W} Hj\left(ct_{ij} - ct_{ij}\right)..\ 2 \quad (9)$$

An algorithm for forecasting financial risk using DAE. Financial and disclosure reports. Prediction of risk. 1) Preprocess the supplied datasets. 2) Extract and pick the necessary characteristics. 3) Categorize the attributes using DAE.θ = $arg_{\theta}m(X,Z)$. Evaluation based on scores using MAE (Equation (8) and MSE (Equation (9). Acquire the anticipated mistakes.

4. RESULT AND DISCUSSION

Data is selected from the KOD CUP 99 (13.3%) dataset and fed into the Deep Auto Encoder (DAE) algorithm. The algorithm is implemented in Python, allowing us to check the detection rate and error detection rate. After completing the training phase, we will analyze the results from the test dataset used during the network training process. Twelve separate experiments were conducted using data from the

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

test DAE algorithm compared with the often-existing algorithm.

The DDoS detection module is thoroughly tested using both strategies, and the experiment outcomes are presented in Figure3. One desirable quality of the DEA algorithm is its low probability of detection for both false positives and negatives. The research compares the results and explores the causes of the risks measured using these two distinct approaches.

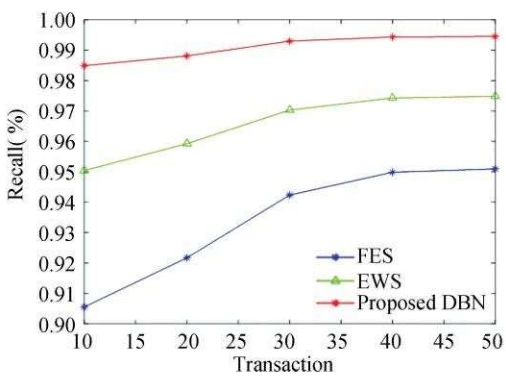

Figure 3: Detection Rate

The research has leveraged various IoT-powered applications to assess, monitor, and investigate potential financial threats. Moving forward, we plan to combine multiple algorithms for comparison, validate them with data, and develop a comprehensive algorithm to identify and calculate financial risk indices. This algorithm will investigate both the safety and threats to financial institutions.

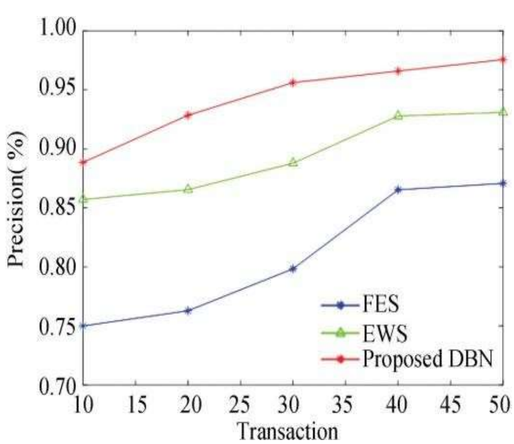

Figure 4: Precision

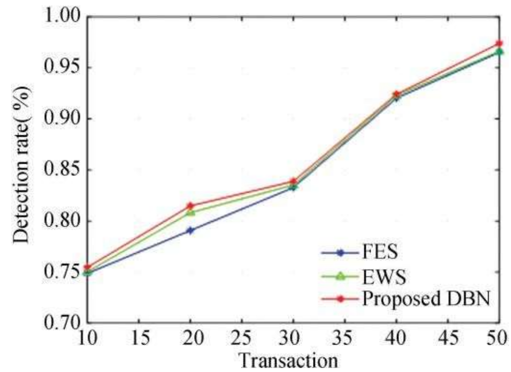

Figure 5: Recall

The network configuration comprises four codes, and we conduct a simulation of the improved Deep Auto Encoder (IDEA) algorithm, comparing its results with those generated by the number of connections in the network. This has led to a rise in the value of the objective function. However, the rate of growth experienced by the IDEA algorithm is significantly slower than that of the basic IDAE algorithm. The optimal reaction generated by the IDEA algorithm is superior to that of the basic IDAE algorithm.

The IDAE algorithm significantly reduces delay jitter and can provide deterministic real- networks. Comparing its results with another algorithm, we find that the DAE algorithm has reduced delay jitter by7.1% compared to the DAE algorithm under stable parameters.

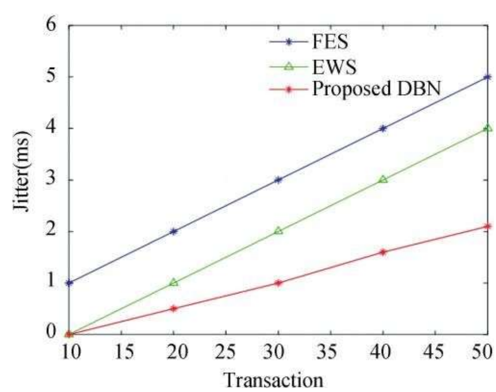

Figure 6: Jitter

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Improving financial risk assessment and management is critical to modern finance, especially in evolving technologies and interconnected global markets. In recent years, researchers and practitioners have increasingly turned to advanced computational techniques, such as deep learning algorithms, to enhance the accuracy and efficiency of financial risk detection and analysis. One promising algorithm is the IDAE (Improved Denoising Autoencoder) algorithm, which builds upon the capabilities of the DAE (Denoising Autoencoder) algorithm to handle large-scale financial data better and improve risk quantification. By leveraging deep learning, IDAE allows for better identification of patterns, anomalies, and trends within financial data, ultimately leading to more effective risk management.

We implemented various configurations to evaluate the effectiveness of the IDAE algorithm in financial risk assessment. We conducted risk measurements across networks of different sizes, ranging from small to large-scale networks. This comprehensive approach allowed us to assess the algorithm's performance across a spectrum of network complexities and sizes. By doing so, we could examine how well the IDAE algorithm adapts to diverse data environments and its ability to efficiently handle varying levels of financial data, providing a robust evaluation of its risk quantification capabilities.

One notable observation from our experiments is the relationship between network size and algorithm performance. As expected, the total time required to execute both DAE and IDAE algorithms increases linearly with the growth in network size. However, we observed an interesting trend regarding running times. While the IDAE algorithm took longer than the DAE algorithm for smaller-scale network configurations, It exhibited shorter running times for slightly larger network topologies. This trend suggests that the scalability and efficiency of the IDAE algorithm improve as the complexity of the network increases, highlighting its suitability for large-scale risk evaluation tasks.

The superior performance of IDAE algorithms in handling larger network topologies can be attributed to their incremental improvements over the DAE algorithm. By incorporating enhanced computational techniques and optimized network configurations, the IDAE algorithm demonstrates improved risk quantification capabilities and faster processing times for complex financial data sets.

Moreover, our findings underscore the advantages of IDAE algorithms in terms of potential risk capture, maximum processing speed, and overall stability. These qualities are crucial for financial institutions and risk managers, who must analyze vast amounts of data accurately and swiftly to make informed decisions and effectively mitigate potential risks.

The training and validation loss curves provide insights into the machine learning model's learning dynamics. The graph below (Fig. 7) depicts the training and validation loss trends over epochs during model training.

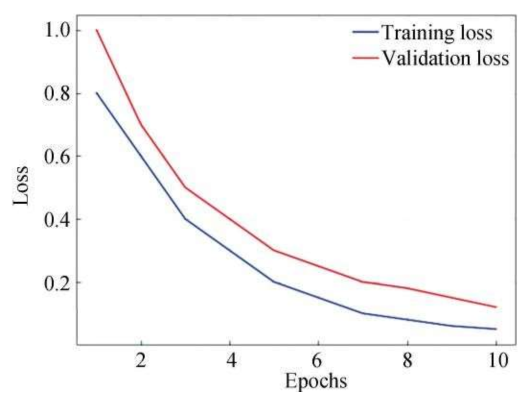

Figure 7: Training and Validation Loss

In addition to performance metrics such as running times and scalability, we also evaluated the effectiveness of the IDAE algorithms in capturing and quantifying various types of financial risks. The algorithm's ability to analyze diverse risk factors and provide meaningful insights is essential for comprehensive risk management strategies.

Furthermore, our experiments highlight the importance of continuous advancements in computational techniques for financial risk assessment. As financial markets evolve and become more interconnected, traditional risk assessment methods may need to be revised to capture emerging risks and market dynamics. Advanced algorithms like the IDAE algorithm offer promising solutions by leveraging machine learning and profound learning principles to uncover hidden patterns and correlations within complex financial data. It is important to note that while IDAE algorithms demonstrate significant improvements over existing approaches, challenges and limitations remain to be addressed. For instance, the computational resources required for running the algorithm on large-scale datasets can be substantial, necessitating efficient hardware infrastructure and optimization techniques. Moreover, as with any algorithmic approach, IDAE algorithms' effectiveness relies on the quality and relevance of the input data. Data preprocessing, including cleaning and feature selection, remains critical in ensuring the algorithms' accuracy and reliability in risk assessment tasks.

5. CONCLUSION

This research emphasizes the considerable influence of broken links and obsolete code on the dependability and security of open-source repositories. These difficulties diminish code efficiency and elevate the danger of vulnerabilities since obsolete or inactive code often lacks prompt updates and fixes. Using an AI autoencoder methodology improves the safety of sensitive data by detecting and signaling possible security vulnerabilities in real-time. The autoencoder's capacity to identify abnormalities in code structures provides a proactive approach to preserving code integrity. This strategy eventually fosters a more secure, efficient, and resilient open-source environment. This article examines the changing financial hazards in the Internet of Things (IoT) ecosystem, using multiple methods to evaluate and quantify risks, detect rogue nodes, and rate overall risk. The results indicate that the DAE algorithm is proficient in quantifying financial risk, providing a reduced false detection rate while preserving accuracy and efficiency in risk evaluation. Using performance-enhancing algorithms yields ideal outcomes in risk management, while the capacity to identify rogue nodes further fortifies security. This thorough study is a significant resource for scholars and practitioners in finance and IoT, high lighting the need for sophisticated algorithms such as DAE to reduce financial risks in linked systems.

REFERENCES:

[1] Xia,Y.,Xu,T.,Wei,M.,etal.(2023).Predictingc hainmanufacturingSMEcreditriskin supply chain finance based on machine learning methods. Sustainability, 15(2), 1087. DOI: 10.3390/su15021087.

[2] Chatterjee, P., Zavadskas, E. K., et al. (2017). Integrated QFD-MCDM framework for green supplier selection. Journal of Cleaner Production, 142(Part 4), 3728–3740. DOI: 10.1016/j.jclepro.2016.10.095.

[3] Li, Y. (2022). Security and risk analysis of the financial industry based on the internet of things. Wireless Communications and Mobile Computing, Article ID 6343468. DOI: 10.1155/2022/6343468.

[4] Mhlanga, D. (2020). Industry 4.0 in finance: The impact of artificial intelligence (AI) on digital financial inclusion. International Journal of Financial Studies, 8(3), 45. DOI: 10.3390/ijfs8030045.

[5] DuHadway, S., Carnovale, S., & Hazen, B. (2019). Understanding risk management for intentional supply chain disruptions: Risk detection, risk mitigation, and risk recovery. Annals of Operations Research, 283, 179–198. DOI: 10.1007/s10479-019-03494-6.

[6] Motoc, M. M. (2022). A proposal for a bankruptcy risk detection model: Adaptation of the Taffler model. Ovidius University Annals Economic Sciences Series.

[7] Khan, A. B., Devi, S., & Devi, K. (2023). An enhanced AES-GCM-based security protocol for securing IoT communication. Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 23(4), 711–719. DOI: 10.17586/2226- 1494-2023-23-4-711-719.

[8] Wang, D. X., Lin, J. B., Cui, P., et al. (2019). A semi-supervised graph attentive network

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for financial fraud detection. 2019 IEEE International Conference on Data Mining (ICDM), 598–607. DOI: 10.1109/ICDM.2019.00072.

[9] Papadakis, S., Garefalakis, A., Lemonakis, C., et al. (2020). Machine learning applications for accounting disclosure and fraud detection. Philadelphia: IGI Global.DOI: 10.4018/978-1-7998-5053-9.

[10] Shi, Y. (2022). A big data technique for internet financial risk control. Mobile Information Systems, Article ID TBD. DOI: TBD.